文章编号: 1005-0523(2006)05-0170-03

基于数据修正的改进型灰色-RBF网络预测

严修红, 许伦辉

(江西理工大学 机电学院, 江西 赣州 341000)

摘要: 针对既呈趋势性, 又呈波动性的时间序列难以预测的问题, 提出了基于失真数据的修正的改进型灰色-RBF网络预测模型及算法. 即用改进型灰色模型提取趋势性因素, 用神经网络处理波动性因素, 另尝试性的提出还须排除异常干扰因素, 即查找和修正序列的异常数据. 并以南昌铁路车站旅客发送量预测为例, 验证了算法的有效性, 收到很好的预测效果.

关键词: 数据修正; 灰色预测模型; 径向基函数神经网络; 排除干扰性.

中图分类号: TP183, U491, O221

文献标识码: A

1 前言

铁路主要车站旅客发送量受多方面因素的影响, 因此从各年的数据来看, 既呈现出呈类似指数增长或下降的长期趋势, 又表现出一定的随机性和波动性. 对于这类时间序列, 文献[1]提出灰色自适应滤波组合优化预测模型, 文献[2]提出一种用于季节性预测的组合灰色人工神经网络模型. 文献[3]提出用灰色模型来进行预测, 然后用神经网络模型对其预测值残差进行修正的灰色组合模型. 本文在以上研究的基础上, 考虑到模型的预测能力不仅是对历史数据的拟合能力, 更重要的是模型的适应能力和预测推广能力, 拟合能力强的适应能力和预测推广能力不一定强, 因为过分的追求拟合将使模型更加复杂, 反而降低了模型的适应能力和推广能力, 因此本文并不着重对历史数据的拟合, 对一些异常数据, 对未来几年预测的参考价值不大, 反而会降低预测结果的准确性和可信性, 我们不去拟合它, 而是去修正它, 同时又充分利用GM模型和神经网络的专长, 提出了基于失真数据的修正的改进型灰色-RBF网络预测模型和算法, 重点讨论了①对数据滑动平均处理, 选取最佳初始值建立GM(1, 1)模型, 提取趋势性因素; ②用RBF网络查找和修

正原始序列的失真数据, 排除预测的干扰因素; ③用GM模型预测值作为RBF网络的输入, 以修正后的实际值作为目标输出对网络进行训练, 训练后仿真得出预测结果, 这样用神经网络处理了随机因素. 大量仿真验证了预测模型及算法的有效性, 收到很好的预测效果.

2 改进的GM(1, 1)模型^{[4]-[5]}

设有变量为 $x^{(0)}$ 的原始非负数据序列, 先对原始数据作滑动平均处理 $x^{(1)} = (1/4)[3x^{(0)}_1 + x^{(0)}_2]$, $x^{(2)} = (1/4)[x^{(0)}_1 + 2x^{(0)}_2 + x^{(0)}_3]$, $x^{(x-1)} = (1/4)[x^{(0)}_{x-2} + 2x^{(0)}_{x-1} + x^{(0)}_x]$, $x^{(n)} = (1/4)[3xx^{(0)}_{x-1} + x^{(0)}_n]$, 最值定解条件取 $x^{(0)}_{k+1} = (x^{(1)}_m - u/a)e^{-a(k-m+1)} + u/a$, $k=1, 2, \dots$ 依次分别选用 $m=1, 2, \dots, n$, 建立GM(1, 1)模型, 得出预测值, 计算预测误差, 通过比较, 选取使预测误差最小的 m 值, 建立最佳预测公式, 从而得出预测值.

3 用RBF网络查找和修正失真数据^[6]

由于某些偶然的因素或特定的情况, 造成了某一年或某几年的数据出现了大的转折, 这样的数据

收稿日期: 2006-01-24

资助项目: 国家自然科学基金项目(60664001), 江西省自然科学基金项目资助(0511030)

作者简介: 严修红(1974-), 男, 江西南康人, 硕士研究生, 研究方向: 交通系统数学建模与预测及仿真.

对未来几年预测的参考价值不大,反而会降低预测结果的准确性和可信性,因此我们必须考虑实际情况,查找出失真(不合理)数据并对其进行修正,从而准确合理的预测.

3.1 失真数据的查找

1) 检验历史各数据是否比前后年份的数据都大(或小),从而判断出该数据是否处于波峰(或波谷),从而把处于波峰或波谷的数据认为是第一种类型的失真数据.

2) 建立一个单输入单输出的 RBF 网络,它的输入输出分别为年份和相应年份的数据.

3) 利用归一化后的历史数据(将上述得到的第一种类型的失真数据排除在外)对 RBF 网络进行训练.

4) 将历史数据中各年份作为输入变量输入到已经完成训练的 RBF 网络,这样就可以得到一组对应的输出值,然后再把这组输出值分别与相应各年份的实际值进行比较,误差超过 5% 的即可认为该年份的数据为失真数据,这是第二种类型的失真数据.

3.2 失真数据的修正

将上述查出的所有失真数据年份作为输入变量输入到已经完成训练的 RBF 网络,这样得到的网络输出就可以作为该年份数据修正值.

4 用 RBF 网络仿真出预测值

建立单输入单输出的 RBF 网络,以改进的 GM(1,1)模型的拟合值或预测值归一化后作为输入,输出结果反归一化后即为最终的预测值.本文采用的是 Matlab 神经网络工具箱中 Solverb 函数编程仿真.

5 基于失真数据的修正的改进型灰色-RBF 网络预测算法步骤

(1) 输入原始数据.

(2) 对原始数据做滑动平均处理,选取最佳初始值建立 GM(1,1)模型,得出拟合值和预测值;提取趋势性因素.

(3) 用 RBF 网络查找和修正原始序列的失真数据,排除预测的干扰因素.

(4) 用 RBF 网络寻求 GM(1,1)预测值和修正后的实际值的复杂的函数关系并仿真出预测值,即处理随机性和波动性.

6 算例与分析

南昌铁路车站旅客发送量见表 1^[7],现以 1996—2003 年的实际人数为训练数据按上述算法预测 2004 年的数据,并对结果进行分析.

(1) 为了便于说明作滑动平均,选最佳定解条件的必要性,笔者以 1996—2003 年的实际人数建立了 3 个 GM(1,1)模型,预测误差结果见表 2,其中 GM1 为原始的 GM 模型,GM2 为对原始数据进行滑动平均后建立的 GM 模型,GM3 是在 GM2 建模基础上再取作为定解条件,平均相对误差指 1996—2003 年的相对误差的绝对值的平均数,预测相对误差指 2004 年的相对误差,且采用了后验差方法进行精度检验,从各项指标来看,对比结果分析可知,作了滑动平均比没做的预测效果好,作了最佳定解条件处理的比没做的要好,特别是预测误差有大幅度的改善.(从 4.86%变为 1.85%)

表 1 南昌车站旅客发送量及其预测值、相对误差

年份	1996	1997	1998	1999	2000	2001	2002	2003	2004
实际人数(万人)	508	666	747	807	795	943	1 013	968	1148
修正后的实际值(万人)	552.3	630.6	747	807	889.4	943	1013	1089.6	1148
改进型 GM 模型预测值(万人)	547.5	675.9	727.1	782.2	841.4	905.2	973.7	1047.5	1126.8
相对误差/%	7.78	1.49	2.66	3.07	5.84	4.01	3.88	8.21	1.85
本文模型最终预测值(万人)	540.1	671.3	731.6	799.1	872.0	947.7	1021.7	1088.3	1140.3
相对误差/%									0.67

(2) 建立单输入单输出 RBF 网络查找和修正失真数据,它的输入输出分别为归一化后的年份和相应的归一化后的年份数据,观察表 1 的实际人数可知,2000 年和 2003 年的人数处于波谷,(事实上,

2003 年的人数少很明显是“非典”的影响,属于特殊情况,对未来预测参考价值不大)属于第一类失真数据,把其他数据归一化后作为训练数据,把所有年份输入后得到一组对应的输出值,然后再把这组

输出值分别与相应各年份的实际值进行比较,误差超过5%的有1996,1997,2000,2003年的数据,相对误差分别为-8.71%,5.31%,-11.87%,-12.57%,从而根据RBF网络输出值分别修正为552.3,630.6,889.4,1089.6.从而得出修正后的实际值(见表1).

表2 南昌铁路车站各种预测方法结果比较

选用模型	平均相对误差(%)	预测误差(%)	后验差系数	小误差概率	精度等级
GM(1,1)1	3.77	4.86	0.2149	1	一级(好)
GM(1,1)2	3.48	2.01	0.2187	1	一级(好)
GM(1,1)3	3.45	1.85	0.2180	1	一级(好)

(3) 建立单输入单输出的RBF网络,把(1)中的改正型GM模型(GM³)的拟合值和预测值(见表2)归一化后的值作为输入,输出的值反归一化后即最终拟合和预测值,训练样本的输入为1996—2003年的改正型GM模型的拟合和预测值归一化后的值,对应的目标值为经修正后的实际人数.仿真后得到的RBF网络中的radbas层的神经元个数为2,2004年的预测值为1140.3,相对误差为0.67%,与实际值非常接近,预测效果相当好.

7 结束语

- 1) 给出了先对数据作滑动平均处理,再选取最佳初始值建立改进GM(1,1)模型方法.
- 2) 对于既呈趋势性,又呈波动性的时间序列的

预测,不仅要用灰色模型提取趋势性因素,神经网络处理波动性因素,本文尝试性的提出还须排除异常干扰因素,即查找和修正序列的异常数据.提出了基于失真数据的修正的改进型灰色-RBF网络预测模型和算法,即基于失真数据的修正的改进型灰色-RBF网络预测模型=趋势预测模型(改进的GM(1,1)模型)+干扰性预测模型(数据修正的RBF网络模型)+周期性、随机性预测模型(RBF网络预测模型),并以南昌铁路车站旅客发送量预测为例,验证了模型及算法的有效性,收到很好的预测效果,为这类时间序列的预测提出了可行性途径.

参考文献:

- [1] 刘兵军,牛东晓.灰色自适应滤波组合优化预测模型及其应用[J].电网技术,1998,22(9):43-45.
- [2] 邢棉.季节性预测的组合灰色神经网络模型研究[J].系统工程理论与实践,2001(1):31-35.
- [3] 王谷,汪洋.神经网络修正灰色残差模型的交通量预测[J].交通标准化,2006(1):76-78.
- [4] 祖恩三.云南GDP的灰色预测和分析[J].经济师,2006(6):272-274.
- [5] 张大海,江世芳,史开泉.灰色预测公式的理论缺陷及改进[J].系统工程理论与实践,2002(8):140-142.
- [6] 陈泽淮,张尧,武志刚.RBF神经网络在中长期负荷预测中的应用[J].电力系统及其自动化学报,2006,18(1):15-19.
- [7] 中国国家统计局,郑京平.中国统计年鉴2005[R].北京:中国统计出版社,2005.

Improved Gray-RBF Network Prediction and Based on Revised Data

YAN Xiu-hong, XU Lun-hui

(Institute of Electromechanical Engineering of Jiangxi University of Science and Technology, Ganzhou 341000, China)

Abstract: Aiming at the problem that the tendency and fluctuated time-series are difficult to forecast, this paper put forward improved gray-RBF network model based on revised data distortion and its algorithms. That is, extracting the trend with the improved gray model, processing the fluctuated factor using radial basis function neural networks, in addition, proposed excluding the unusual interference factors, namely to find out and identify abnormal data of the sequence. and the forecasting of passenger transmission volume in Nanchang Railway Station verified the validity of the model and algorithm, received a good predictive effect.

Key words: revised data; gray forecasting model; radial basis function neural network; excluding the unusual interference factors